客体类别差异调节视听跨通道冲突中反应水平的感觉主导效应*

周衡[1,2,3] 王爱君[1] 袁祥勇[2,3] 蒋毅[2,3]

([1]苏州大学心理学系, 心理与行为科学研究中心, 苏州 215123)

([2]中国科学院心理研究所脑与认知科学国家重点实验室, 北京 100101)

([3]中国科学院大学心理学系, 北京 100049)

摘　要 感觉主导效应是大脑面临多种感觉通道信息时, 优先加工某种感觉通道信息的现象。认知加工水平假说认为感觉主导效应的产生是由于不同认知加工水平所决定, 即早期知觉加工水平表现为视觉主导, 晚期反应加工水平表现为听觉主导。然而, 现有的研究并未关注认知加工水平中处于早期和晚期之间的中间加工水平如何影响感觉主导效应。研究操纵了位于中间加工水平的客体类别差异, 采用 2-1 mapping(映射)范式通过 3 个实验考察介于早期知觉和晚期反应水平之间的客体类别表征如何影响跨通道冲突感觉主导效应。实验 1 结果发现, 客体类别差异能够调节反应水平的感觉主导效应。类别差异小时, 表现为视觉主导; 类别差异大时, 表现为听觉主导。实验 2 结果说明此效应的产生与视觉刺激不同加工深度无关, 证实了此效应的产生特异于视觉通道。实验 3 通过经颅直流电刺激技术抑制类别加工脑区, 即左侧颞前叶。结果发现, 反应水平的听觉主导效应消失。研究表明, 认知加工水平中的客体类别表征这一中间加工水平对感觉主导效应产生了调节, 完善了认知加工水平假说对于跨通道冲突感觉主导效应的解释。

关键词 客体类别, 视觉主导效应, 听觉主导效应, 反应前水平, 反应水平

收稿日期：2024-06-06

*江苏省社会科学基金 24JYB003; 苏州大学人文社科青年交叉团队

周衡和王爱君同为第一作者

通信作者: 王爱君, E-mail: ajwang@suda.edu.cn; 蒋毅, E-mail: yijiang@psych.ac.cn

1 引言

多感觉整合是将多个感觉通道的信息整合为一个连贯的有意义的知觉过程，以帮助我们更快更好地知觉外界信息，对我们的生存发展具有重要意义(Stein & Stanford, 2008; Tang et al., 2016)。当我们面对多个感觉通道的不同信息时，我们的大脑并不会给予每个感觉通道以相同的权重，这就导致某些感觉通道的信息优先被加工，从而产生感觉主导效应(Callan et al., 2015; Hirst et al., 2018; Zhou et al., 2022)。感觉主导效应中最为常见的是视觉主导效应和听觉主导效应。如视觉主导效应发现，当视觉闪光和听觉声音同步呈现时被试倾向于只对视觉闪光进行反应而忽略听觉声音(Colavita, 1974; Wang et al., 2021)。听觉主导效应发现，当视觉闪光与听觉声音个数不匹配时，被试对视觉闪光个数的判断会受听觉声音的影响(Shams et al., 2000; 王爱君 等, 2020)。然而，感觉主导效应如何产生以及受何种因素制约是该领域的重要问题之一(Keil, 2020; Hirst et al., 2020)。

研究者提出不同的理论解释感觉主导效应的机制。注意定向假说认为，注意指向的感觉通道信息优先加工(Amedi et al., 2017)，但其无法解释经典的听觉主导效应。通道适宜性假说认为，最适合特定任务的感觉通道驱动感知(Hirst et al., 2020)，但其忽略了自上而下的先验知识或期望的影响(Wang et al., 2019)。基于数学的贝叶斯因果推断模型结合感觉信号相对可靠性、数量先验预期及对共同来源的先验预期来确定感觉主导效应的产生(Odegaard et al., 2016; Wozny et al., 2008)，尽管此模型能特异性地很好地解释听觉主导效应，但对视觉主导效应的解释有限。上述理论假说更多强调外部信号的影响，而忽视了内部加工水平的作用。被试在不同实验范式下会产生不同的感觉主导效应，如同一组被试在 McGurk 范式下表现为视觉主导效应，在声音诱发闪光错觉范式下表现为听觉主导效应 (Butera et al., 2023)，这表明尽管实验范式改变了感觉主导效应的方向，但个体内部的信息加工过程或许是感觉主导效应发生的重要前提(Robinson et al., 2016)。因而，研究者从认知加工本身的角度解释感觉主导效应的发生，进而提出了加工水平假说。加工水平假说认为，感觉主导效应由认知加工水平决定，早期知觉加工水平表现为视觉主导效应，晚期反应加工水平表现为听觉主导效应(Chen & Zhou, 2013; Li et al., 2019)。

早期研究者将认知加工水平划分为早期知觉加工水平和晚期反应加工水平(王勇慧 等, 2006)。陈安涛等人(2010)认为在早期知觉水平和晚期反应水平之间还存在着思维加工的中间模块加工水平。一些研究者意识到语义加工等中间加工水平的作用，但是承认难以分离知觉和语义加工，因此将认知加工水平划分为含知觉和语义加工的反应前水平以及反应水平

(Chen & Zhou 2013)。然而，现有关于认知加工水平下的感觉主导效应主要聚焦于早期知觉加工水平(Koppen et al., 2009)和晚期反应加工水平(Kato & Konishi, 2006; Mayer et al., 2009), 而尚未关注介于两者之间的中间加工水平如何影响反应水平的感觉主导效应。此外, 以往认知加工水平下的感觉主导效应研究采用 2-1 mapping 范式, 使用政治人物和明星人物两种类别的视觉图片和对应听觉发音名字作为刺激材料, 在该范式下, 刺激-反应映射构成三种条件, 一致条件, 图片和声音一致；反应前不一致条件, 图片和声音不匹配但对应同一按键；反应不一致条件, 图片和声音不匹配且对应不同按键。反应前不一致条件和一致条件之间的差异反映了反应前水平冲突, 而反应不一致条件和反应前不一致条件之间的差异反映了反应水平冲突。研究者通过此范式将认知加工水平划分为反应前水平和反应水平, 结果发现在反应前水平, 注意听觉时视觉干扰物造成的冲突显著大于注意视觉时听觉干扰物的冲突, 即产生了视觉主导效应, 与之相对的是, 反应水平表现为听觉主导效应(Chen & Zhou, 2013; Li et al., 2019)。然而该研究使用的政治人物和明星人物分属于不同社会身份, 类别差异较大，是需要个体后天长期学习和经验的结果。并且面孔加工具有特异性, 与动物等其他类别刺激的加工存在差异(Martin & Chao, 2001)。前人研究使用动物的图片和声音作为刺激材料, 发现视觉干扰物的干扰效应强于听觉干扰物, 即出现视觉主导效应(Molholm et al., 2004)。这表明先前研究发现不同的感觉主导效应的方向可能受到客体类别差异的影响。

客体类别加工发生在介于早期知觉加工水平和晚期反应加工水平之间的中间加工水平(Bi et al., 2016; Martin, 2007)。以往研究表明, 客体识别在很大程度上取决于对客体形状的处理(Li et al., 2024; Morgenstern et al., 2024), 当客体的结构相似性较高时，识别反应时较短、正确率较高；而当客体的结构相似性较低时，被试的正确率越低。动物类客体在视觉外观和功能属性方面是高度相似的, 结构差异较小, 而工具类客体由于视觉和功能更加多样化, 结构差异较大(Wiggett et al., 2009)。Vogler 和 Titchener (2011)使用动物和工具类客体作为实验材料, 发现动物类客体被识别的速度和准确性显著高于工具类客体。与此一致, Laws (2000)发现, 与工具类客体相比, 被试对动物类客体的视觉和听觉加工能力更强。鉴于不同客体类别的加工存在差异, 本研究通过客体类别表征差异这一认知加工中的中间水平加工探究内部表征如何影响感觉主导效应。

本研究采用 2-1 mapping 范式, 操纵客体类别的差异, 考察客体类别对跨通道冲突感觉主导效应的影响。实验 1 操纵客体类别的差异程度探究其是否以及如何影响感觉主导效应。实验 1a~c 的刺激材料分别为类别差异较小的动物类客体(实验 1a：野生动物 vs. 家养动物)、

类别差异中等的工具类客体(实验 1b：交通工具 vs. 家用工具)和类别差异较大的动物和乐器类客体(实验 1c：动物 vs. 乐器)。由于客体类别加工在早期知觉加工水平之后，因此客体类别加工可能会影响晚期反应加工水平的输出，实验 1 假设客体类别影响反应水平的感觉主导效应。鉴于实验 1 视觉刺激以图片形式呈现，其表征速度快于听觉通道呈现的名字。因此，实验 2 将视觉刺激由图片呈现改为文字呈现，以此来探讨客体类别对反应水平感觉主导效应的影响是否特异于视觉通道而非不同层次的视觉加工。尽管图片和文字加工速度不同，但均能在反应水平之前完成加工，因此实验 2 假设客体类别的呈现方式不会受到影响反应水平的感觉主导效应。鉴于左侧颞前叶负责客体类别加工(Binney et al., 2018; Diez et al., 2017; Wong & Gallate, 2012)，因此实验 3 在实验 1 的基础上进一步采用经颅直流电刺激技术(transcranial direct current stimulation, tDCS)抑制左侧颞前叶，因果性地探究客体类别对感觉主导效应的影响。实验 3 假设神经调控左侧颞前叶，反应前水平不会受到影响，而反应水平的感觉主导效应发生改变。

2 实验 1：客体类别差异对视听跨通道冲突中反应水平感觉主导效应的调节作用

2.1 方法

2.1.1 被试

为了观察到稳定的跨通道冲突效应，本研究使用 G*Power 3.1 (Faul et al., 2007, 2009)的配对样本 t 检验对样本量进行估算，参考 Chen 和 Zhou (2013)的研究中跨通道冲突效应 Cohen's d = 0.70，设置 α err prob = 0.05, power (1 - β err prob) = 0.8，输出 total sample size = 19，因此实验样本量应不低于 19 人。为了确保效应量在统计水平上足够大，实验 1a 共招募了 31 名大学生，其中 1 名被试被剔除(由于对目标反应的正确率< 80%而被剔除，由于任务较为简单，其余被试的人均正确率在 90%以上)，剩余 30 名被试(男性 10 名，女性 20 名，平均年龄 20.97 岁，SD = 2.04)；实验 1b 共招募 32 名被试，其中 2 名被试被剔除，剩余 30 名被试(男性 12 名，女性 18 名，平均年龄 20.43 岁，SD = 1.79)；实验 1c 共招募 30 名被试(男性 12 名，女性 18 名，平均年龄 20.33 岁，SD = 1.92)，所有被试的视力或矫正视力正常，且均不了解本实验的目的。所有被试均为右利手，实验前签署了知情同意书且在实验过后获得相应的报酬。本研究依据赫尔辛基宣言进行，并经苏州大学批准。

2.1.2 实验仪器和材料

本实验在安静、微暗的实验室内进行。实验程序由 Matlab 2014b 和 Psychtoolbox 工具包

编写，搭载在 Windows 7 操作系统上。实验刺激在 27 英寸的 LCD 液晶显示器上呈现(刷新率 60 Hz，分辨率 1920 × 1080)，实验过程中被试坐在距显示器约 60 cm 处，显示器屏幕上的背景色设为灰色(RGB: 127, 127, 127)。位于灰色屏幕中央的“+”(视角 1.5º × 1.5º)作为注视点在整个实验过程中始终呈现。

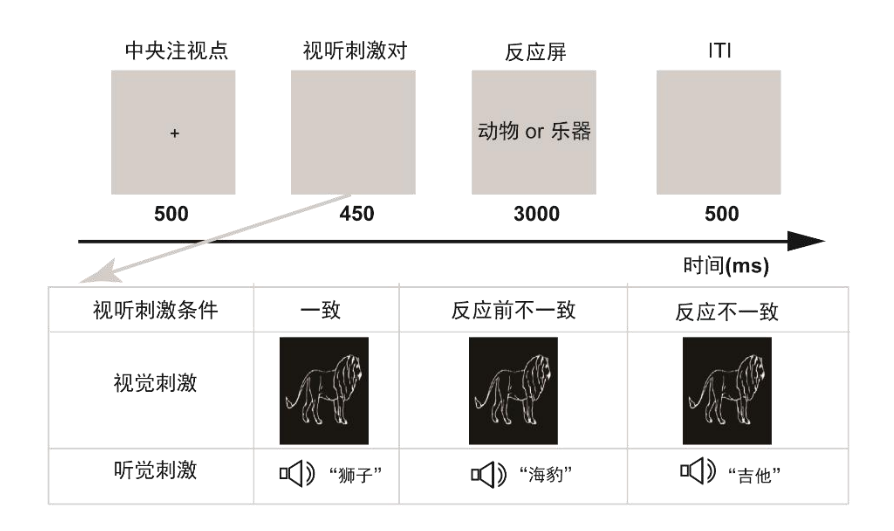
每个实验的视觉刺激材料均是从经过一系列评定的标准轮廓图库中选取的 (Snodgrass & Vanderwart, 1980)，实验前经过 30 名被试 (16 名女性，平均年龄 20.94, *SD* = 1.76) 的命名和熟悉性评分，最终在命名正确率和熟悉性超过 90%的图片中挑选出 6 张作为实验材料。在实验 1a 中，图片分别为野生动物(大象、狮子、海豹)和家养动物(母鸡、鸭子、猫咪)，在实验 1b 中，图片分别为交通工具(汽车、飞机、轮船)和家用工具(熨斗、水壶、椅子)，在实验 1c 中，图片分别为动物(狮子、大象、海豹)和乐器(吉他、钢琴、长笛)。图片材料(视角 5º × 6º)在屏幕中央呈现 450 ms，听觉刺激为 6 张图片对应的汉语名字，通过头戴式耳机(ATH-WS99)与视觉刺激同时呈现，音量在 65 dB，呈现时间为 450 ms。

2.1.3 实验设计和流程

所有实验设计均为 2 (注意通道：注意视觉、注意听觉) × 3 (一致性：一致、反应前不一致、反应不一致)的被试内设计，实验 1a 的任务为判断所注意通道呈现的是野生动物还是家养动物，实验 1b 的任务为判断所注意通道呈现的是交通工具还是家用工具，实验 1c 的任务为判断所注意通道呈现的是动物还是乐器。对于被试来说，在整个实验中，所有的图片和名字都是需要回应的潜在目标。图片和名字构成了三种一致性条件，一致条件：图片和名字是同一个客体(如实验 1c 的看到狮子和听到“狮子”)，反应前不一致条件：图片和名字不对应同一客体，但对应同一按键(如实验 1c 的看到狮子和听到“海豹”)，反应不一致条件：图片和名字对应不同按键(如实验 1c 的看到狮子和听到“吉他”)。

注意通道以 block 设计呈现，三种一致性试次在 block 里随机混合，在每个 block 中，被试被要求注意视觉或听觉，同时忽略来自其他通道的刺激。在整个实验过程中，被试被要求盯着屏幕中央的十字加号。500 ms 的中央注视点结束后，视觉刺激和听觉刺激同时呈现 450 ms，被试有 3000 ms 的时间进行反应，在进入下一个试次前有 500 ms 的试次间隔。实验共有 24 个 block，每个 block 下有 12 个试次, 3 种一致性试次随机呈现。注意视觉和注意听觉的 block 交替出现，顺序在被试间平衡，在每一个 block 开始前有一个 2000 ms 的指导语，告知被试注意视觉或注意听觉。实验共有 6 种实验条件，每种实验条件有 48 个试次，总共 288 个试次，为了保持被试在实验中精力充沛，实验过程中设计了 3 次休息，正式实验时长大约

30 分钟，正式实验前有大约 3 分钟的练习。



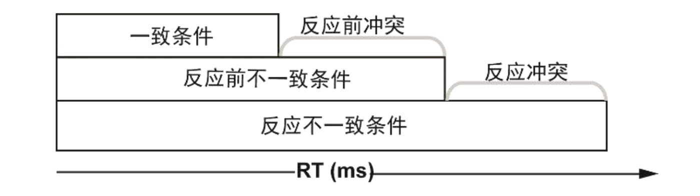


图 1 实验 1c 的刺激实例及实验流程图。视听刺激构成三种条件，一致条件：视听刺激完全一致；反应前不一致条件：视听刺激不一致但对应按键一致；反应不一致条件：视听刺激对应按键不一致。反应前冲突 = 反应前不一致条件的反应时 − 一致条件的反应时；反应冲突 = 反应不一致条件的反应时 − 反应前不一致条件的反应时。

2.1.4 数据分析

首先，由于实验被试的正确率均高于 90%，本研究不再对错误试次进行分析，仅对正确试次进行分析，并将与平均反应时相差 3 个 *SDs* 的试次排除在分析之外 (原始反应时数据见下表)。其次，本研究使用各视听条件反应时的差值来区分反应前水平和反应水平的冲突效应，反应前不一致条件与一致条件的反应时差值即为反应前水平的冲突效应，反应不一致条件与反应前不一致条件的反应时差值即为反应水平的冲突效应。最后，对比注意视觉时听觉干扰物导致的冲突效应与注意听觉时视觉干扰物导致的冲突效应，当视觉冲突显著大于听觉冲突时，即为产生视觉主导效应；当听觉冲突显著大于视觉冲突时，即为产生听觉主导效应。研究报告η_p^2 和 Cohen’s *d* 作为 *F* 检验和 *t* 检验的效应量指标。

表 1 各实验中被试的原始反应时平均值和标准误(ms)

实验名称	注意视觉			注意听觉		
	一致	反应前不一致	反应不一致	一致	反应前不一致	反应不一致
实验 1a	494±20	542±29	539±29	756±25	850±29	884±32
实验 1b	522±32	571±40	602±45	787±26	873±339	907±34
实验 1c	496±30	529±37	562±45	777±27	862±30	867±30
实验 2	503±27	535±33	559±36	791±24	874±26	880±29
实验 3-伪	520±22	522±25	591±28	704±34	761±38	807±33
实验 3-阴	510±24	539±27	580±38	679±26	734±29	773±27

2.2 结果

实验 1 操纵客体类别的差异程度探究其对反应前水平和反应水平感觉主导效应的影响。实验 1a~c 分别为类别差异较小的动物类客体(实验 1a：野生动物 vs. 家养动物), 类别差异中等的工具类客体(实验 1b：交通工具 vs. 家用工具)以及类别差异较大的动物和乐器类客体(实验 1c：动物 vs. 乐器)。在实验 1 中, 视听刺激构成三种条件：一致条件, 反应前不一致条件以及反应不一致条件。反应前不一致条件与一致条件的反应时差值即为反应前水平的冲突效应, 反应不一致条件与反应前不一致条件的反应时差值即为反应水平的冲突效应。对比注意视觉时听觉干扰物造成的冲突与注意听觉时视觉干扰物造成的冲突, 探究反应前水平和反应水平的感觉主导效应以及客体类别对其的影响。

在实验 1a 中(图 2A), 对于反应前水平, 视觉冲突(94)显著大于听觉冲突(48), $t(29) = 3.78$, $p < 0.001$, Cohen's $d = 0.69$, 95% CI = [21.13, 71.06]。对于反应水平, 视觉冲突(34)显著大于听觉冲突(-3), $t(29) = 2.81$, $p = 0.009$, Cohen's $d = 0.51$, 95% CI = [10.11, 64.13]。该结果表明不管是反应前水平还是反应水平, 视觉冲突均显著高于听觉冲突, 即存在视觉主导效应。

在实验 1b 中(图 2B), 对于反应前水平, 视觉冲突(87)显著大于听觉冲突(49), $t(29) = 2.58$, $p = 0.015$, Cohen's $d = 0.47$, 95% CI = [7.88, 68.09]。对于反应水平, 视觉冲突与听觉冲突没有显著差异, $t < 1$, $BF_{10} = 0.197$, 有较中等的证据支持零假设。该结果表明在反应前水平, 视觉冲突显著高于听觉冲突, 即存在视觉主导效应, 而在反应水平, 视觉冲突与听觉冲突相当, 即不存在感觉主导效应。

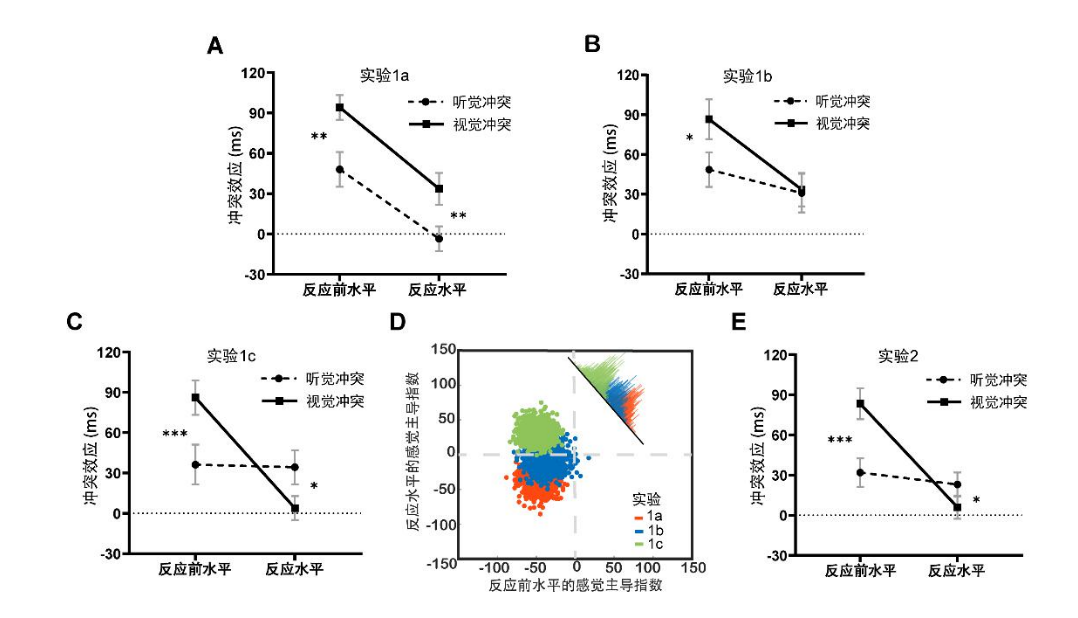


图 2 实验 1 和 2 的结果图。A～C 分别为实验 1a～c 的结果图；D 为实验 1 每个实验的感觉主导指数(听觉冲突与视觉冲突之差)再抽样得到的样本均值分布图；E 为实验 2 的结果图。注：$^{***}p < 0.001$, $^{**}p < 0.01$ 和$^{*}p < 0.05$, Bonferroni corrected。误差棒表示标准误。

在实验 1c 中(图 2C)，对于反应前水平，视觉冲突(86)显著大于听觉冲突(36), $t(29) = 4.33$, $p < 0.001$, Cohen’s $d = 0.79$, 95% CI = [26.29, 73.32]。对于反应水平，听觉冲突(34)显著大于视觉冲突(4), $t(29) = 2.53$, $p = 0.017$, Cohen’s $d = 0.46$, 95% CI = [5.82, 55.03]。该结果表明在反应前水平,视觉冲突显著高于听觉冲突，即存在视觉主导效应，而在反应水平，听觉冲突显著高于视觉冲突，即存在听觉主导效应。

为了进一步比较实验 1 三个子实验中感觉主导效应的强度，本研究计算了每个子实验的感觉主导指数，方法是用听觉冲突减去视觉冲突(正值表示存在听觉主导效应，负值表示存在视觉主导效应)。将感觉主导指数进行 3 (实验：1a、1b、1c) × 2 (冲突水平：反应前水平、反应水平)的混合方差分析，结果发现，实验主效应显著, $F(2, 87) = 4.94$, $p = 0.009$, $\eta_p^2 = 0.10$。冲突水平主效应显著, $F(1, 87) = 10.40$, $p = 0.002$, $\eta_p^2 = 0.11$。二者交互作用有显著的趋势, $F(2, 87) = 2.61$, $p = 0.080$, $\eta_p^2 = 0.06$, $BF_{10} = 3.21$, $BF_{10} > 3$ 有较强证据支持非零假设。为了探究潜在交互作用的影响，我们分别对反应前水平和反应水平的感觉主导指数进行了简单效应分析。对于反应前水平，单因素方差分析的结果表明三个子实验不存在显著差异, $F < 1$。与 0 进行单样本 t 检验得, $t(89) = 6.05$, $p < 0.001$, Cohen’s $d = 0.64$, 95% CI = [29.98, 59.28]。该结果表明

在反应前水平上，始终产生视觉主导效应，客体类别并未影响反应前水平的感觉主导效应。

对于反应水平，单因素方差分析结果表明实验主效应显著，$F(2, 87) = 6.00$, $p = 0.004$, $\eta_p^2 = 0.12$。基于 Bonferroni correction 的事后分析结果发现，实验 1c 的感觉主导指数显著大于实验 1a, $t(59) = 3.46$, $p = 0.002$, Cohen's $d = 0.89$, 95% CI = [21.05, 114.03]。而实验 1b 与实验 1a、1c 的差值均不存在显著差异，$t_1(59) = 1.77$, $p = 0.24$; $t_2(59) = 1.69$, $p = 0.28$。与 0 进行单样本 t 检验得，实验 1b 与 0 不存在显著差异，$t < 1$，说明不存在感觉主导效应，实验 1a 显著小于 0, $t(29) = 2.81$, $p = 0.009$, Cohen's $d = 0.51$, 95% CI = [10.11, 64.13]，说明存在视觉主导效应；实验 1c 显著大于 0, $t(29) = 2.53$, $p = 0.017$, Cohen's $d = 0.46$, 95% CI = [5.82, 55.03]，说明存在听觉主导效应。该结果表明在反应水平上，感觉主导效应受到客体类别的调节，客体类别差异较小时，表现为视觉主导效应，反之，则为听觉主导效应。

为了更加直观观察以及验证各实验的结果，本研究通过标准的 bootstrap 方法绘制了三个子实验中感觉主导指数 1000 次重采样的结果分布(图 2D)，横坐标表示反应前水平的感觉主导指数，纵坐标表示反应水平的感觉主导指数。三个实验均主要分布在横坐标 0 的左侧；然而实验 1a 主要分布在纵坐标 0 下方，实验 1b 分布在纵坐标 0 附近，实验 1c 主要分布在纵坐标 0 上方。该结果同样说明，客体类别调节反应水平的感觉主导效应。

3 实验 2：排除视觉刺激呈现方式差异的影响

尽管实验 1 的结果表明客体类别影响反应水平的感觉主导效应，然而，实验 1 的视觉刺激均以图片的形式呈现，图片刺激能表现出客体的外观结构，更早通达知觉表征(Baddeley& Hitch, 2017; Potter & Fox, 2009; Wiggett et al., 2009)，而听觉刺激则最早通达语义表征。根据感觉-语义模型(Higdon et al., 2024; Roberts et al., 2022)，图片能表现出更为细微的感觉代码，例如各种姿势的猫的图片，却只对应"猫"这一语义代码。图片刺激相比于文字或声音刺激，在加工上更占优势(Hockley & Bancroft, 2011)，因此视觉刺激的呈现方式可能会对实验 1 的结果造成影响。基于此，实验 2 在实验 1c 的基础上，将视觉刺激由图片呈现转换为文字呈现，以此来探讨客体类别的呈现方式对反应水平感觉主导效应的影响。根据经典认知理论(Sternberg, 2000)，语义加工处于知觉水平和反应水平之间，因此，无论图片还是文字均能在反应水平之前完成语义加工。实验假设即使视觉刺激的呈现形式由图片转变为文字，反应水平的感觉主导效应仍然不会受到影响。

3.1 方法

3.1.1 被试

实验 2 共招募了 35 名大学生, 其中 1 名被试被剔除(标准与实验 1 相同), 剩余 34 名被试(男性 15 名, 女性 19 名, 平均年龄 21 岁, *SD* = 1.92), 所有被试的视力或矫正视力正常, 且均不了解本实验的目的。所有被试均为右利手, 实验前签署了知情同意书且在实验过后获得相应的报酬。本研究依据赫尔辛基宣言进行, 并经苏州大学批准。

3.1.2 实验仪器和材料

实验仪器同实验 1c, 视觉刺激材料与实验 1c 有一点不同, 由图片刺激转变为文字刺激, 字体是 SimSun, 视角 5º × 6º。

3.1.3 实验设计和流程

同实验 1c。

3.1.4 数据分析

同实验 1c。

3.2 结果

在实验 2 中(图 2E), 对于反应前水平, 视觉冲突(83)显著大于听觉冲突(32), $t(33) = 5.45$, $p < 0.001$, Cohen's $d = 0.93$, 95% CI = [32.16, 70.54]。对于反应水平, 听觉冲突(23)显著大于视觉冲突(6), $t(33) = 2.26$, $p = 0.031$, Cohen's $d = 0.39$, 95% CI = [1.67, 32.44]。该结果表明在反应前水平, 在反应前水平, 视觉冲突显著高于听觉冲突, 即存在视觉主导效应, 而在反应水平, 听觉冲突显著高于视觉冲突, 即存在听觉主导效应。

不仅如此, 将实验 1c 和实验 2 在反应前水平和反应水平的视觉冲突和听觉冲突分别进行独立样本 *t* 检验, 均未发现显著差异, $ts < 1$, $BF_{10} < 0.33$, 贝叶斯因子的结果表明有较为中等的证据支持零假设。该结果说明视觉刺激由文字呈现时, 反应水平的感觉主导效应模式仍未发生改变, 表明视觉刺激的呈现方式不会影响反应水平的感觉主导效应。

4. 实验 3：因果性验证客体加工对视听跨通道冲突中反应水平感觉主导效应的影响

采用 tDCS 技术从因果层面验证客体类别对反应水平感觉主导效应的影响。以往研究表明, 左侧颞前叶是客体类别加工的重要脑区(Binney et al., 2018; Diez et al., 2017)。因此实验 3 在实验 1c 的基础上, 对左侧颞前叶施加阴极刺激, 观察其行为结果。实验 3 假设客体类别加工受到抑制, 反应前水平仍表现为视觉主导效应, 而反应水平的听觉主导效应将发生改变。

4.1 方法

4.1.1 被试

实验 3 共招募了 20 名大学生 (男性 9 名, 女性 11 名, 平均年龄 21.50 岁, $SD = 2.63$), 所有被试的视力或矫正视力正常, 且均不了解本实验的目的。所有被试均为右利手, 实验前签署了知情同意书且在实验过后获得相应的报酬。本研究依据赫尔辛基宣言进行, 并经苏州大学批准。

4.1.2 实验仪器和材料

实验仪器采用美国的 SOTERIX 公司生产的 TES2001 经颅直流电刺激仪, 刺激电极片包裹在大小为 $5 \times 7 = 35$ cm^2 的海绵片中, 电流为 2 mA, 密度为 0.06 mA/cm^2, 按照国际 EEG 10-20 系统的标准和相应的磁共振成像研究, 使用电极帽进行定位, 对于左侧前颞叶, tDCS 的阴极电极放于 T7 和 FT7 之间, 阳极电极放在对侧眶上区, 这是针对左侧前颞叶皮层的 tDCS 研究常用的刺激位置(Akbiyik et al., 2020; Binney et al., 2018)。实验中的正式刺激大小是 2 mA, 时长为 10 分钟, 刺激开始和结束的时候有 30 秒的淡入淡出时间, 伪刺激大小为 2 mA, 时长为 30 秒, 具有相同的淡入淡出时间, 电极放置时长与正式刺激相同。

其余与实验 1c 相同。

4.1.3 实验设计和流程

被试在实验开始之前先进行 10 min 的微弱电刺激, 其余实验流程与实验 1c 相同。被试首先被随机分配到不同刺激组并接受不同的 tDCS 刺激, 顺序在被试间平衡, 每个被试需在不同日内完成 2 次 tDCS 刺激, 每次 tDCS 刺激后完成一次实验任务。

4.1.4 数据分析

同实验 1c。

4.2 结果

伪刺激条件下(图 3C), 对于反应前水平, 视觉冲突(57)显著大于听觉冲突(2), $t(19) = 4.34$, $p < 0.001$, Cohen's $d = 0.97$, 95% CI = [28.56, 81.71]。对于反应水平, 听觉冲突(69)显著大于视觉冲突(46), $t(19) = 2.13$, $p = 0.047$, Cohen's $d = 0.48$, 95% CI = [0.39, 45.91]。该结果与实验 1c 一致, 反应前水平表现出视觉主导效应, 反应水平表现出听觉主导效应。

在阴极刺激条件下(图 3D), 对于反应前水平, 视觉冲突(55)显著大于听觉冲突(29), $t(19) = 2.66$, $p = 0.015$, Cohen's $d = 0.60$, 95% CI = [5.48, 45.94]。对于反应水平, 视觉冲突与听觉冲突并不存在显著差异, $t < 1$。该结果表明经过阴极刺激后, 反应前水平仍表现出视觉主导效应, 而反应水平的感觉主导效应则发生了改变。

为了进一步探究阴极刺激是否对反应前水平的感觉主导效应的大小造成了改变, 实验 3

将伪刺激条件和阴极刺激条件的感觉主导指数进行对比，发现反应前水平的感觉主导指数不存在显著差异, $t(19) = 1.84$, $p = 0.082$, Cohen’s $d = 0.41$, 95% CI = [-4.13, 62.97], $BF_{10} = 0.95$, 即阴极刺激并没有改变反应前水平的视觉主导的大小。以上结果表明，在左侧前颞叶施加阴极刺激后，反应前水平的视觉主导效应未受影响，而反应水平的听觉主导效应消失。

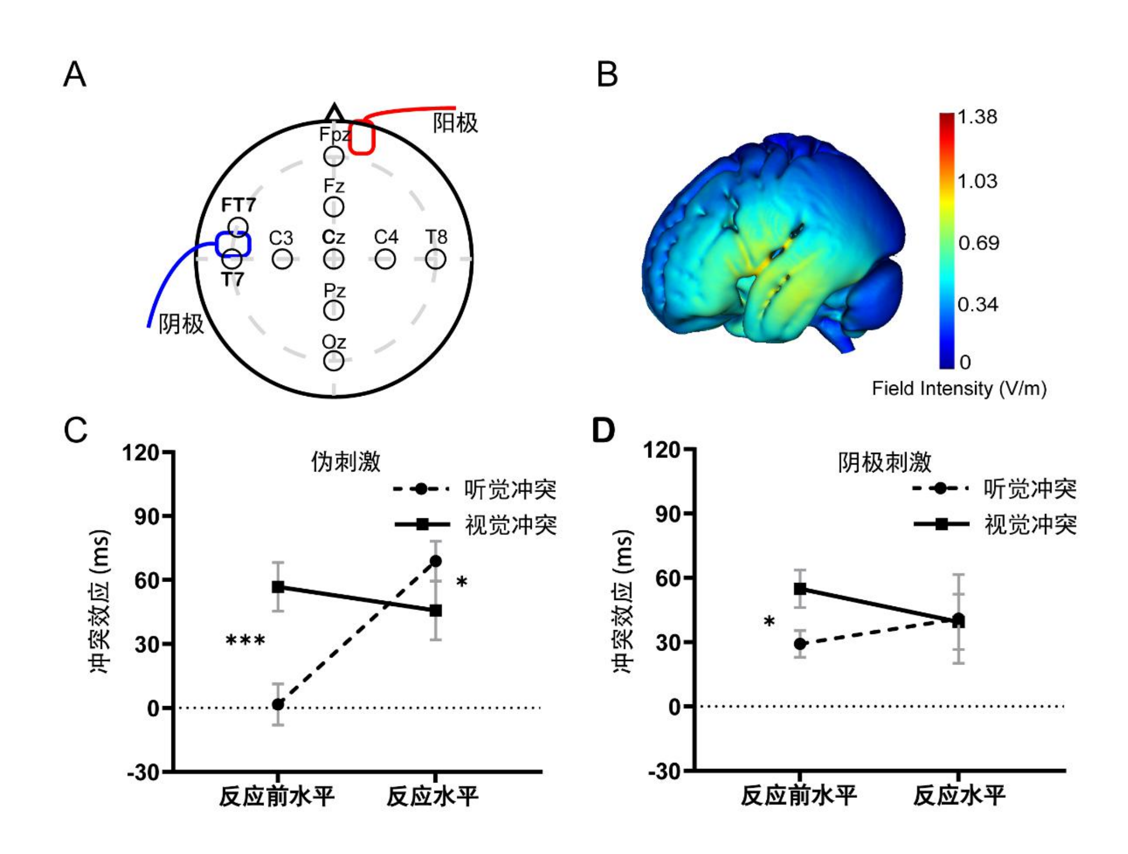


图 3 A 为左侧前颞叶皮层在国际 EEG 10-20 系统标准的定位示意图，位于 T7 和 FT7 之间；B 为 HD-explore 软件模拟左侧前颞叶皮层在标准脑模型的电场模型，C 为伪刺激条件下的实验结果；D 为阴极刺激条件下的实验结果。注：$^{***}p < 0.001$, $^{*}p < 0.05$, Bonferroni corrected。误差棒表示标准误。

5 讨论

本研究使用 2-1 mapping 范式，通过 3 个实验，探究客体类别如何影响反应水平的感觉主导效应。首先，实验 1 发现客体类别差异调节反应水平的感觉主导效应，具体表现为当客体类别差异较小(实验 1a)时，反应水平产生视觉主导效应；当客体类别差异中等(实验 1b)时，反应水平不表现感觉主导效应；而当客体类别差异较大(实验 1c)时，反应水平产生听觉主导效应。不仅如此，实验 2 在实验 1c 的基础上改变视觉刺激的呈现方式，结果发现反应水平的感觉主导效应与实验 1c 类似，表明客体类别对反应水平的感觉主导效应的影响与呈现方式

无关。实验 3 在客体类别加工的重要脑区左侧颞前叶施加阴极 tDCS 刺激后，发现反应水平的感觉主导效应消失，进一步验证了客体类别对反应水平感觉主导效应的影响。

本研究发现反应前水平视觉占据主导地位，这与先前研究结果一致(Chen & Zhou, 2013; Koppen et al., 2009)，这可能是早期加工过程中视听不对称的体现。先前研究发现，当存在跨通道冲突时，视觉干扰物的干扰效应要显著强于听觉干扰物的干扰效应(康冠兰，罗霄骁，2020; Donohue et al., 2013)。定向注意理论认为，视觉刺激的加工相较于自动化加工的听觉刺激而言，个体警觉性更低，需要更多的注意资源(Dietze & Poth, 2023)。在本研究中，视觉刺激和听觉刺激的输入是同时的，个体在反应前水平的信息加工包含知觉和语义水平的加工，听觉刺激被自动化加工，占据的注意资源较少，而视觉刺激占据了大量的注意资源，因此导致了视觉刺激的优先加工。

本研究发现中间加工水平对反应水平的感觉主导效应起到调控作用。类别差异较大时，反应水平表现为听觉主导效应；类别差异较小时，表现为视觉主导效应。扩散启动模型(spreading priming model, Schubert, 2021)或许能对其进行解释，视觉干扰物激活的额外无关的语义表征干扰了听觉目标的判断，尤其是在语义相关性较强的情况下，视觉干扰物激活了更多额外的语义表征，因此产生更强的冲突效应。本研究的结果也与进化观点一致，即注意资源的分配受生物体当前需求的影响，个体根据环境线索做出适宜性权衡，如避免危险和选择性认知协调(Haselton & Buss, 2000)。当一条蛇伪装在草丛中时，如果它发出声音，就更容易分辨出来。在当前的研究中，当跨通道的客体类别差异较大时，可以通过听觉通道进行更为准确区分。本研究与 Chen 和 Zhou (2013)的发现具有不同之处，该研究采用政治人物和明星人物的面孔和名字作为实验材料，发现反应前水平表现为视觉主导效应，反应水平表现为听觉主导效应。本研究实验 1c(类别差异较大)与该研究的结果一致，可能原因在于人脸具有特殊性(Martin & Chao, 2001; Ubaldi & Fairhall, 2021)，同时政治人物和明星人物分属两个社会类别，即使有很多相似的特征，例如两个眼睛一个鼻子，人类还是能很快对其进行区分，而熟悉的人脸更是如此，甚至研究者发现有专门存放个人信息(如特征、态度、身份地位等)的独特的高维的个体信息空间(Castello et al., 2021)。

尽管实验 1 表明客体类别影响反应水平的感觉主导效应，然而并不能排除自下而上的刺激呈现方式的影响。实验 1 的视觉刺激由图片呈现，听觉刺激则是客体的中文名字。图片更早通达知觉表征，口语与文字更早通达语义表征(Baddeley& Hitch, 2017; Potter & Fox, 2009; Wiggett et al., 2009)，加工速度的不匹配或许影响实验 1 的结果。先前研究表明，对于图片而

言，与理解意义相比，理解文字名称代码的时间相对较晚，而对于单词而言，理解意义的时间相对较晚(Dirani & Pylkkänen, 2023)。在分类任务中，图片被分类的速度通常快于文字(Dirani & Pylkkänen, 2023)，而在命名任务中，文字被命名的速度快于图片(Nieznański, 2020)。然而，尽管不同呈现方式的加工时间或有不同，但类别表征的加工均在反应水平之前实现，因此客体类别的呈现方式不会对反应水平的感觉主导效应产生影响，实验 2 的结果支持这一假设。

最后，本研究对左侧颞前叶施加伪刺激和阴极 tDCS 刺激(Akbiyik et al., 2020; Binney et al., 2016; Mesulam et al., 2013)，左侧颞前叶已被广泛证明在客体命名和识别中发挥重要作用，语义性痴呆患者的该部位损伤，导致概念知识受损，从而导致泛化错误(Lambon Ralph & Patterson, 2008)和基于语义关联的记忆错觉(Binney et al., 2018; Diez et al., 2017)。颞前叶在听觉通道加工中起重要作用，其活动受抑制或许影响听觉加工，进而影响感觉主导效应。实验 3 发现，无论采用伪刺激还是阴极刺激，反应前水平始终保持着稳定的视觉主导效应，且视觉主导效应的大小不受 tDCS 刺激的影响，而施加阴极刺激减弱其皮层兴奋性后，反应水平的感觉主导效应发生改变。抑制颞前叶后，反应前水平的感觉主导效应不变说明更可能是客体类别加工受抑制，而非听觉加工受抑制，然而，反应前水平的贝叶斯因子的结果较弱地支持零假设，因此，以上结果并不能完全排除听觉加工受抑制的可能性，未来研究需进一步排除听觉加工的影响。

研究者提出了各种理论解释感觉主导效应的发生机制。例如，定向注意假说从注意的角度解释了感觉主导效应的方向，即拥有更多注意资源的感觉通道是占据主导地位的(Amedi et al., 2017)。通道适宜性假说认为，感觉通道驱动的感知最适合于特定的任务(Hirst et al., 2020)，如视觉在空间任务中占据主导地位，而听觉在时间任务中占据主导地位。然而，以上假说均不能解释本研究反应前水平和反应水平出现不同感觉主导效应的现象。不仅如此，本研究的视听刺激均能被清晰知觉，且不涉及先验预期等信息，因此，本研究的感觉主导现象也不能通过贝叶斯因果推断模型解释。本研究的结果支持感觉主导效应与人类本身的认知加工水平有关，感觉主导效应在不同的认知加工水平上表现不同(Chen & Zhou, 2013; Li et al., 2019)。

本研究首次以客体类别差异为切入点观察到中间加工水平对反应水平的感觉主导效应起到调控作用。尽管研究发现，感觉主导效应在不同的加工水平有不同的表现，如早期知觉加工水平的视觉占主导和晚期反应加工水平听觉占主导(Chen & Zhou, 2013; Li et al., 2019)，

介于两者之间的中间加工水平负责思维、语义等内部表征(陈安涛 等, 2010; Velasquez et al., 2021), 对人类高级认知和社会交互具有重要作用。然而, 研究者对中间加工水平在多感觉整合中所起的作用知之甚少。本文借助客体类别这一内部表征, 探究中间加工水平如何影响感觉主导效应, 进一步对加工水平假说进行完善, 即早期知觉加工水平视觉占主导, 晚期反应加工水平的感觉主导受到中间加工水平的灵活调节, 随内部表征的不同, 晚期反应加工水平输出的感觉主导不同。后续研究探究多感觉整合问题时, 应将负责内部表征加工的中间加工水平考虑在内。

6 结论

通过操纵客体类别的差异程度, 中间加工水平可以调节反应水平的感觉主导效应, 具体表现为客体类别差异较小时, 反应水平表现为视觉主导, 客体类别差异较大时, 反应水平表现为听觉主导, 这种行为模式与视觉刺激呈现方式无关, 但受到左侧颞前叶活动的影响。

参考文献

Akbiyik, S., Goksun, T., & Balci, F. (2020). Cathodal tDCS stimulation of left anterior temporal lobe eliminates cross-category color discrimination response time advantage. *Behavioural Brain Research, 391*, 112682.

Amedi, A., Hofstetter, S., Maidenbaum, S., & Heimler, B. (2017). Task selectivity as a comprehensive principle for brain organization. *Trends in Cognitive Sciences, 21*(5), 307–310.

Baddeley, A. D., & Hitch, G. J. (2017). Is the levels of processing effect language-limited? *Journal of Memory and Language, 92*, 1–13.

Bi, Y. C., Wang, X. Y., & Caramazza, A. (2016). Object domain and modality in the ventral visual pathway. *Trends in Cognitive Sciences, 20*(4), 282–290.

Binney, R. J., Ashaie, S. A., Zuckerman, B. M., Hung, J. Y., & Reilly, J. (2018). Frontotemporal stimulation modulates semantically-guided visual search during confrontation naming: a combined tDCS and eye tracking investigation. *Brain and Language, 180*, 14–23.

Butera, I. M., Stevenson, R. A., Gifford, R. H., & Wallace, M. T. (2023). Visually biased perception in cochlear implant users: a study of the McGurk and Sound-induced flash illusions. *Trends in Hearing, 27*, 23312165221076681.

Callan, A., Callan, D., & Ando, H. (2015). An fmri study of the ventriloquism effect. *Cerebral cortex, 25*(11), 4248–4258.

Castello, M. V., Haxby, J. V., & Gobbini, M. I. (2021). Shared neural codes for visual and semantic information about familiar faces in a common representational space. *Proceedings of the National Academy of Sciences, 118*(45). e2110474118.

Chen, A. T., Liu, Q., Chen, C. M., & Li, H. (2010). The effect of Intermediate processing level on response conflict in Stroop-like tasks. *Journal of Psychological Science, 33*(3), 569–572.

[陈安涛, 刘强, 陈昌明, 李红. (2010). 中间模块加工水平对反应冲突的影响及其机制. *心理科学, 33*(3), 569–572.]

Chen, Q., & Zhou, X. L. (2013). Vision dominates at the preresponse level and audition dominates at the response

level in cross-modal interaction: behavioral and neural evidence. *Journal of Neuroscience, 33*(17), 7109–7121.

Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics, 16*(2), 409–412.

Dietze, N., & Poth, C. H. (2023). Vision rivals audition in alerting humans for fast action. *Acta Psychologica, 238*, 103991.

Dirani, J., & Pylkkänen, L. (2023). The time course of cross-modal representations of conceptual categories. *NeuroImage, 277*, 120254.

Donohue, S. E., Todisco, A. E., & Woldorff, M. G. (2013). The rapid distraction of attentional resources toward the source of incongruent stimulus input during multisensory conflict. *Journal of Cognitive Neuroscience, 25*(4), 623–635.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160.

Haselton, M. G., & Buss, D. M. (2000). Error management theory: a new perspective on biases in cross-sex mind reading. *Journal of personality and social psychology, 78*(1), 81–91.

Higdon, K. F., Neath, I., Surprenant, A. M., & Ensor, T. M. (2024). Distinctiveness, not dual coding, explains the picture-superiority effect. *Quarterly Journal of Experimental Psychology.* https://doi.org/10.1177/17470218241235520

Hirst, R. J., Cragg, L., & Allen, H. A. (2018). Vision dominates audition in adults but not children: a meta-analysis of the Colavita effect. *Neuroscience & Biobehavioral Reviews, 94*, 286–301.

Hirst, R. J., McGovern, D. P., Setti, A., Shams, L., & Newell, F. N. (2020). What you see is what you hear: twenty years of research using the sound-induced flash illusion. *Neuroscience & Biobehavioral Reviews, 118*, 759–774.

Hutmacher, F. (2019). Why is there so much more research on vision than on any other sensory modality? *Frontiers in Psychology, 10*, 2246.

Kang, G. L., & Luo, X. X. (2020). The integration and conflict control in audiovisual processing. *Journal of Psychological Science, 43*(5), 1072–1078.

[康冠兰, 罗霄骁. (2020). 视听跨通道信息的整合与冲突控制. *心理科学, 43*(5), 1072–1078.]

Kato, M., & Konishi, Y. (2006). Auditory dominance in the error correction process: a synchronized tapping study. *Brain Research, 1084*(1), 115–122.

Keil, J. (2020). Double flash illusions: current findings and future directions. *Frontiers in Neuroscience, 14*, 298–304.

Koppen, C., Levitan, C. A., & Spence, C. (2009). A signal detection study of the Colavita visual dominance effect. *Experimental Brain Research, 196*(3), 353–360.

Lambon Ralph, M. A., & Patterson, K. (2008). Generalization and differentiation in semantic memory: insights from semantic dementia. *Annals of the New York Academy of Sciences, 1124*, 61–76.

Laws, K. R. (2000). Category-specific naming errors in normal subjects: the influence of evolution and experience. *Brain and Language, 75*(1), 123–133.

Li, Z., Gu, R., Qi, M., Cen, J., Zhang, S., Gu, J., Zeng, X., & Chen, Q. (2019). Loss of vision dominance at the preresponse level in tinnitus patients: preliminary behavioral evidence. *Frontiers in Neuroscience, 13*, 482.

Li, X., Cai, S. Z., Chen, Y., Tian, X. M., & Wang, A. J. (2024). Enhancement of visual dominance effects at the response level in children with attention-deficit/hyperactivity disorder. *Journal of Experimental Child Psychology. 242*, 105897.

Marks. L. E. (2004). Cross-modal interactions in speeded classification. In: Calvert. G. A., Spence. C., Stein. B. E. (Eds.), *The handbook of multisensory processes.* (pp. 85–105). Cambridge: MIT.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58*, 25–45.

Martin, A. & Chao, L. (2001). Semantic memory and the brain: structure and processes. *Current Opinion in Neurobiology, 11*(2), 194–201.

Mayer, A. R., Franco, A. R., Canive, J., & Harrington, D. L. (2009). The effects of stimulus modality and frequency of stimulus presentation on cross-modal distraction. *Cerebral cortex, 19*(5), 993–1007.

Mesulam, M. M., Wieneke, C., Hurley, R., Rademaker, A., Thompson, C. K., Weintraub, S., & Rogalski, E. J. (2013). Words and objects at the tip of the left temporal lobe in primary progressive aphasia. *Brain, 136*(2), 601–618.

Molholm, S., Ritter, W., Javitt, D. C., & Foxe, J. J. (2004). Multisensory visual-auditory object recognition in humans: A high-density electrical mapping study. *Cerebral Cortex, 14*(4), 452–465.

Morgenstern, Y., Storrs, K. R., Schmidt, F., Hartmann, F., Tiedemann, H., Wagemans, J., & Fleming, R. W. (2024). High-level aftereffects reveal the role of statistical features in visual shape encoding. *Current Biology, 34*(5), 1098–1106.

Nieznański, M. (2020). Levels-of-processing effects on context and target recollection for words and pictures. *Acta Psychologica. 209*, 103127.

Odegaard, B., Wozny, D. R., & Shams, L. (2016). The effects of selective and divided attention on sensory precision and integration. *Neurosciences Letters, 614*, 24–28.

Potter, M. C., & Fox, L. F. (2009). Detecting and remembering simultaneous pictures in a rapid serial visual presentation. *Journal of Experimental Psychology-Human Perception and Pertypyance, 35*(1), 28–38.

Roberts, B. R.T., MacLeod, C. M., & Fernandes, M. A. (2023). Symbol superiority: why $ is better remembered than ‘dollar’. *Cognition, 238*. 105435.

Robinson, C. W., Chandra, M., & Sinnett, S. (2006). Existence of competing modality dominances. *Attention, Perception, & Psychophysics, 78*(4), 1104–1114.

Schubert, E. (2021). Creativity is optimal novelty and maximal positive affect: a new definition based on the spreading priming model. *Frontiers in Neuroscience, 15*, 612379.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature, 408*(6814), 788–788.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 picture: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 6*(2), 174–215.

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience, 9*(4), 255–266.

Tang, X. Y., Wu, J. L., & Shen, Y. (2016). The interactions of multisensory integration with endogenous and exogenous attention. *Neuroscience & Biobehavioral Reviews, 61*, 208–224.

Ubaldi, S., & Fairhall, S. L. (2021). fMRI response to automatic and purposeful familiar-face processing in perceptual and non-perceptual cortical regions. *Journal of Neurophysiology, 125*(4), 1058–1067.

Velasquez, A. G., Gazzaley, A., Toyoda, H., Ziegler, D. A., & Morsella, E. (2021). The generation of involuntary mental imagery in an ecologically-valid task. *Frontiers in Psychology. 12*, 759685.

Vogler, J. N., & Titchener, K. (2011). Cross-modal conflicts in object recognition: determining the influence of object category. *Experimental Brain Research, 214*, 597–605.

Wang, A. J., Huang, J., Lu, F. F. He, J. Y., Tang, X. Y., & Zhang, M. (2020). Sound-induced flash illusion in multisensory integration. *Advances in Psychological Science, 28*(10), 1662–1677.

[王爱君, 黄杰, 陆菲菲, 何嘉滢, 唐晓雨, 张明. (2020). 多感觉整合中的声音诱发闪光错觉效应. *心理科学*

进展, 28(10), 1662–1677.]

Wang, A., Zhou, H., Zhang, F., Sang, H., Yu, W., Tang, X., Zhang, T., & Zhang, M. (2021). Repetition suppression in visual and auditory modalities affect the sound-induced flash illusion. *Perception, 50*(6), 489–507.

Wang, A. J., Sang, H. B., He, J. Y., Sava-Segal, C., Tang, X. Y., & Zhang, M. (2019). Effects of cognitive expectation on sound-induced flash illusion. *Perception, 48*(12), 1214–1234.

Wang, Y. H., Wang, Y. F., & Zhou, X. L. (2006). Conflict Control at different periods of processing in children with two subtypes of ADHD. *Acta Psychologica Sinica, 38*(2), 181–188.

[王勇慧, 王玉凤, 周晓林. (2006). 注意缺陷多动障碍儿童在不同加工阶段的干扰控制. *心理学报, 38*(2), 181–188.]

Wiggett, A. T., Pritchard, R. C., & Downing, P. E. (2009). Animate and inanimate objects in human visual cortex: evidence for task-independent category effects. *Neuropsychologia, 47*(14), 3111–3117.

Wong, C., & Gallate, J. (2012). The function of the anterior temporal lobe: a review of the empirical evidence. *Brain Research, 1449*, 94–116.

Wozny, D. R., Beierholm, U. R., & Shams, L. (2008). Human trimodal perception follows optimal statistical inference. *Journal of Vision, 8*(3), 24.

Zhou, H., Li, S., Huang, J., Yang, J., Wang, A., & Zhang, M. (2022). Sound-induced flash illusions at different spatial locations were predicted by personality traits. *Attention, Perception, & Psychophysics, 85*(2), 463–473.

Object category differences regulate the sensory dominance of the response level in an audiovisual cross-modal conflict

ZHOU Heng[1,2,3], WANG Ai-Jun[1], YUAN Xiang-Yong[2,3], JIANG Yi[2,3]

([1]*Department of Psychology, Research Center for Psychology and Behavioral Sciences, Soochow University, Suzhou* 215123, *China*)

([2]*State Key Laboratory of Brain and Cognitive Science, Institute of Psychology, Chinese Academy of Sciences, Beijing*100101, *China*)

([3]*Department of Psychology, University of Chinese Academy of Sciences, Beijing*100049, *China*)

Abstract

Sensory dominance is a phenomenon in which the brain selectively processes specific sensory information when presented with multisensory inputs, thereby enhancing human perception of external stimuli. Previous studies have discussed sensory dominance at perceptual and response levels. However, how the intermediate processing level between perceptual and response levels affects sensory dominance remains unknown. Therefore, the present study adopted the cross-modal 2-1 mapping paradigm, and object categories were manipulated through three studies to investigate the role of the intermediate processing level on sensory dominance in a cross-modal conflict.

In this paradigm, based on key mapping, cognitive processing levels can be defined into the preresponse level (including perceptual and semantic levels) and the response level. The difference between the audiovisual incongruent condition and the audiovisual congruent condition is called the conflict effect, and sensory dominance can be obtained by comparing the conflict effect of attention to vision and auditory. Experiment 1 manipulated the degree of difference in object categories to explore its impact on sensory dominance. Experiments 1a – 1c involved animal objects (small differences), tool objects (moderate differences), and animal and musical instrument objects (large differences). A total of 30 participants were recruited for each experiment. Visual pictures reached perceptual representation early, whereas auditory sounds reached semantic representation early. Therefore, Experiment 2 (34 participants) changed visual pictures into visual words on the basis of Experiment 1c to explore the effects of the visual presentation way of object categories on sensory dominance. In Experiment 3 (20 participants), transcranial direct current stimulation (tDCS) was used on the left anterior temporal lobe, an important brain region responsible for processing object categories to study casually the effects of object category on the sensory dominance of the response level further.

The results of Experiment 1 showed that, regardless of the difference in object categories, the conflict effect of attention to auditory at the preresponse level was significantly greater than that of attention to vision, that is, visual dominance. However, visual dominance at the response level appeared when the object category difference was small (Experiment 1a). Moreover, no sensory dominance was observed when the object category difference was moderate (Experiment 1b), and auditory dominance appeared when the object category difference was large (Experiment 1c). The results of Experiment 2 and Experiment 1c were consistent, that is, auditory dominance, indicating that this behavior pattern was not affected by the bottom-up visual presentation way. The results of Experiment 3 showed that under the cathodal tDCS condition, the preresponse level still showed visual dominance. However, the response level no longer showed sensory dominance. This result showed the effects of object categories on the sensory dominance of the response level from the causal level.

The mechanism of sensory dominance is still under investigation. The present study was the first to find that object categories affected the sensory dominance of the response level. From the perspective of cognitive processing level, the intermediate processing level played a regulating role in the sensory dominance of the response level. This finding can enrich the explanatory theory of sensory dominance and can provide a new perspective for the study of sensory dominance in a cross-modal conflict.

Keywords object categories, visual dominance, auditory dominance, preresponse level, response level